智力科学源于生命科学无法解决的重要问题，哪是什么？还需要科学逻辑：一阶逻辑陷阱是造成人工智能测试失败和无效的原因。一阶逻辑无法找出层出不穷的错误。

如何突破用逻辑研究逻辑、用测量研究测量的困境？本杰明·富兰克林研究电时，还没量子物理和凝聚态物理，也不可能准确定义电。牛顿说：I do not d efine time, space, place and motion, as being well known to all。科学基于自然语言，但有更严谨的规范。

我的科学逻辑和智力科学新理论事先准确分析、预测了欧盟Human Brain project和美国BRAIN initiative的失败和问题。对人工智能和语言模型的分析正在验证中。

没忘帝国兴衰和过度扩张？人工智能跟逻辑简单、可靠的传统计算机技术如计算器、编译器、互联网有本质区别。计算机已扩张到力不从心领域。吸毒适合炒作股票？

万事开头难，要整理出思路不容易。很多人因此放弃思考，求助于人工智能。但人工智能可靠吗，真比吸毒有效？很多人类工作领域都需要高阶逻辑智力，人工智能却没有。

人工智能无法理解自然语言语义和科学理论，就无法理解生理和疾病。人类医生依赖不可靠人工智能，自身诊断能力退化，测试医学人工智能就更困难，这种熵趋势太可怕。

人工智能改进很快趋缓，自动编程也会如此。高昂投入成本下庞氏链无法持久，软件工程师薪水会向其它行业趋近，看不见的手会起作用。硅谷和斯坦福模式不得不转型。

Nvidia、OpenAI爆发前很久，我就批评人工智能问题、倡导新逻辑和新智力研究。是他们想碾压我的科学理论。黄仁勋先生还指望五年内实现通用人工智能AGI、万亿人形机器人市场和人工智能医生？我不研究光刻机和芯片制造，不参与Nvidia布局Intel之类活动。

面对庞杂训练大数据，计算机无法有效区分真假、对错，无法化解别人留下的语言陷阱和情绪陷阱，还会漏洞百出地胡编乱造。这些跟word-embed ded vector space等等技术和构造性数学无法正确表达自然语言语义有关。即使有些具体错误被纠正，结构上root cause无法纠正，更多、更严重问题会产生。更好处理器和训练方法也无法解决。

几十年前就能用计算机模拟核试验。也有实用EDA自动设计电路。EDA逻辑简单、测试容易，使用不可靠人工智能也无妨。但太多复杂逻辑问题测试困难，纠错更难或不可能。数字孪生和工业自动化一直在成长，但增速和取代人力幅度低于预期。人工智能出路在于更有效人机耦合，研究相应智力结构、机制和接口需要新的科学逻辑和智力科学。

说计算机神经网络是物理学在误人子弟。大学研究和教育怎么办？下列重要问题需研究：

1. 通用人工智能AGI是否可能？研究切入点：荷尔蒙、思维模式切换和演变、如何克服熵趋势。荷尔蒙会引发fight-or-flight等浅层快捷反应模式、影响理性和价值观；而克服熵趋势是思维模式演变和提升智力的关键。

2. 人工智能包括人形机器人的智力上限在哪里，需要什么人机耦合结构、机制和接口？

3. 计算机语言模型会制造什么样的诡辩和误导、导致什么样的智力退化？

统计方法对人工智能测试无效。《人形机器人的产业陷阱》附录一有我对Google Search上AI overview一些简单测试结果，可说明计算

机语言模型不懂自然语言语义，更不懂科学理论。这些涉及“知之为知之，不知为不知”等等高阶逻辑抽象问题。所以需要进一步研究逻辑主义、直觉主义、形式主义数学学派的本质区别和各自缺陷以及如何修正、改进。

现在学术界、产业界所用人工智能测试方法完全错误。我把计算机围棋、语言模型、人形机器人当作人工智能技术初级、中级、高级实验工具，但不看好其商业应用。

Richard Sutton 说大语言模型是死路一条，但他倡导的松鼠和强化学习虽适用于有明确控制边界的围棋，却缺乏足够抽象和逻辑能力、无法达到一般人类智力。所以计算机围棋只能做低级实验工具。手工标识则会堵死智力进化途径。人形机器人可进一步验证生命智力和机器智力的区别、思维模式的切换和演变以及智力的进化和退化。

Judea Pearl 的因果理论缺乏严谨、有效的概念和方法，无法准确设计人工智能测试，也无法分析用逻辑研究逻辑、用测量研究测量会造成比哥德尔语句更重要、更广适问题。

古典含义上的 Critique 需要超越、探索边界并形成更好抽象。我不使用逻辑上有问题的 critical thinking 和实用主义教育家杜威使用的 reflective thinking。Critique 一词更严谨，可以跟边界、临界点、奇点等等概念构成有效的逻辑体系。

未对现有脑/神经和智力研究作系统、审慎评估，轻率使用炒作语言把计算机神经网络当物理学是明显的智力退化。上述这些问题，斯坦福大学“以人为本”人工智能研究院和吴蔡神经研究所知道吗，怎么教学生？所以说：前沿研究性大学的教育和研究急需改革。

计算机自然语言处理的进展，比我想象的更差。以下是几个层次的情况：

1. 把语音转化成文字。这是较低的听觉层次。我不研究，所以曾以为已基本解决，实际上还差很远。

2. 翻译。我曾以为大部分已解决，实际上也没有。举个简单例子：俄语 Слышу малиновый звон，слышу 是 hear，малиновый 是 raspberry 或颜色 crimson，звон 是铃或铃声。Google Translate 译成 I hear the crimson ringing，当然错误。我俄语非常差，也知道应该是 raspberry。虽然 Google 可以很快纠正已公布的具体问题，但结构上的 root cause 无法纠正。涉及重要的抽象和高阶逻辑，错误会很多。还常有些低级错误，如某些部分内容在翻译中消失了，人类很容易看出来。所以人类智力和计算机智力明显不同。

3. 根据现有信息回答问题。人工智能往往比人类死记硬背还差，不但会胡编乱造，简单换个方式提问往往就会答错、前言不搭后语。

4. 解决新问题。对少数问题有些价值如：计算机围棋有明确边界，蛋白质结构也是高度受限的局部问题。但人类未来可能找到比人工智能更好的理论解法。而绝大部分问题缺乏很好的边界控制，人工智能就不行了。事实上，语言模型技术已露败相。